Sickle Gene

ITS ORIGIN AND DIFFUSION FROM WEST AFRICA

J. GREGORY MEARS and HERBERT M. LACHMAN, *Department of Medicine, Albert Einstein College of Medicine, Bronx, New York 10461*
RAYMOND CABANNES, *Centre Hospitalier Univérsitaire d'Abidjan, Côte d'Ivoire*
KOSSI P. E. AMEGNIZIN, *Faculté de Médecine, Université du Benin, Lome, Togo*
DOMINIQUE LABIE, *Institut de Pathologie Moléculaire, Institut National de la Santé et de la Recherche Médicale U15, Paris, France*
RONALD L. NAGEL, *Department of Medicine, Albert Einstein College of Medicine, Bronx, New York 10461*

ABSTRACT Linked DNA polymorphisms can be used to study the evolution of structural gene mutations. Both the β^S-($\beta^{6\,Glu\rightarrow Val}$) and β^C-($\beta^{6\,Glu\rightarrow Lys}$) genes are common in West Africa. We have analyzed their linkage to a polymorphic Hpa 1 site appearing 3′ to the β-globin gene locus in selected populations from West Africa. A large reservoir of β^A-genes linked to 13-kilobase Hpa 1 fragments with a frequency of 17–18% has been identified. In addition, the β^S- and β^C-genes in Togo are found to be tightly linked to the 13-kilobase Hpa 1 fragment, whereas 72% of the β^S-genes in the Ivory Coast reside on the 7.6-kilobase Hpa 1 fragment. These studies are consistent with the selection and expansion of two different chromosomes bearing β^S-genes in at least two physically close, but ethnically separate regions of West Africa, with subsequent diffusion to North, Equatorial, and East Africa.

INTRODUCTION

The sickle gene (β^S)[1] mutation is prevalent in areas of the world where malaria has been endemic by providing a selective advantage for the heterozygote (1). It is not known whether the point mutation responsible for the β^S-gene arose once, and then diffused to other regions, or whether it arose independently in diverse regions under the same selective pressure. Additionally, the hemoglobin C mutation (β^C) is found commonly in a more circumscribed region of West Africa. Evidence suggests that the β^C-gene mutation may not be in equilibrium, and that the homozygote is more fit than the heterozygote or normals (2). We report studies of a polymorphic restriction endonuclease recognition site associated with the β-globin gene locus that suggest that the β^S-gene was selected for at least twice in West Africa, whereas the β^C-gene was selected for once. The data are compatible with the diffusion of two chromosomes bearing β^S-genes to other parts of the world.

Restriction endonuclease analysis of the human β-globin gene and its neighboring regions has resulted in a wealth of information concerning its normal fine structure (3, 4) as well as the alterations responsible for some of the hereditary disorders of hemoglobin production (3–5). Recently, Kan and Dozy described a restriction endonuclease polymorphism associated with the β-globin gene locus (6). The normal β-gene (β^A) in Caucasian Americans resides on a 7.6-kilobase (kb) DNA fragment when analyzed by digestion with the restriction endonuclease Hpa 1 and Southern blotting (6, 7). In Black Americans, the β^S-gene is found on a 7.6-kb fragment approximately one-third of the time, whereas two-thirds of the β^S-genes are found on a larger 13-kb fragment (5–7) (Fig. 1). The β^A-gene in Black Americans is commonly on a 7.6-kb fragment, though 3% (6) to 8–9% (7, 8) are found on a 13-kb fragment. Because the β^S-gene in the United States is derived from diverse origins in Africa, the Hpa 1 polymorphism can be used to trace the origin and diffusion of the sickle mutation. Observations of small numbers of Africans suggest that β^S-13-kb Hpa 1 fragment linkage is common in West Africa and rare in East Africa (7, 9).

We have studied the linkage of the Hpa 1 polymorphism to the β^A-, β^S-, and β^C-genes in selected populations from West Africa. Togo was chosen as one study population because evidence from mapping of β^C-gene frequency clines suggests that the β^C-gene arose in the

A preliminary report of this work was submitted in abstract form to the Annual Meeting of the American Society for Clinical Investigation, San Francisco, 25–27 April, 1981.
Dr. Mears is a Hartford Fellow.
Received for publication 16 March 1981 and in revised form 15 May 1981.

[1] *Abbreviations used in this paper*: β^S, sickle globin gene; β^C, hemoglobin C gene; β^A, normal β-globin gene; kb, kilobase.

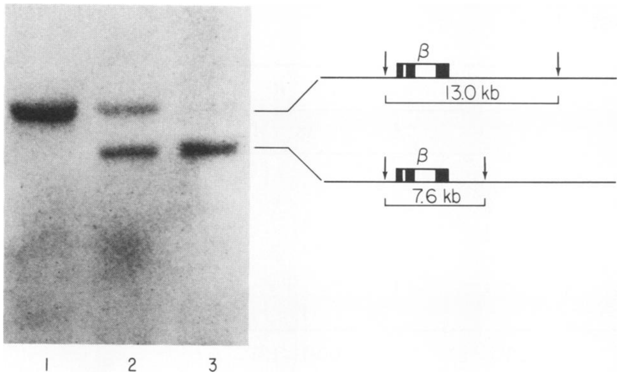

FIGURE 1 Hpa 1 analysis of β-globin genes. Channel 1 contains DNA from a Togolese SS patient. Only a 13-kb Hpa 1 fragment is found. Channels 2 and 3 contain DNA from Togolese AA individuals. Sample 2 contains a 13-kb and a 7.6-kb Hpa 1 fragment, whereas sample 3 contains only a 7.6-kb Hpa 1 fragment. The scheme to the right represents the Hpa 1 polymorphism. The arrows indicate the Hpa 1 recognition sites surrounding the β-globin gene which is represented by rectangles (black, exon; white, intron).

region (2, 10), and because the β^S-gene frequency is estimated to be as high as 16% (11). The ethnic groups represented include the Kabye, Ewe, and Mina. The Ivory Coast is a nearby region of West Africa peopled by unrelated ethnic groups, but which also harbor the β^S-gene in high frequency (11). Ethnic groups represented include the Mande, Kru, Akan, and Voltaique.

The results indicate that, contrary to Kan and Dozy's proposal (9), the Hpa 1-β^S gene linkage is not homogeneous in West Africa. Instead the results suggest that two different chromosomes bearing the β^S-gene were selected for within the region. The appearance of the β^S-gene on the 7.6-kb Hpa 1 fragment in other parts of Africa and the world does not require recurrent independent mutations (9) and may be explained by gene flow from West Africa.

METHODS

105 unrelated individuals were studied, 51 from Togo and 54 from the Ivory Coast. Analysis of the tribal origin of the subjects' parents demonstrated that intertribal marriage was limited.

Peripheral blood samples were obtained, with informed consent, from the study population. Hemoglobin electrophoreses were performed to identify hemoglobin type. Leukocytes pellets were prepared and digested with Proteinase K in the presence of 0.5% sodium dodecyl sulfate as previously described (6). The samples were extracted with phenol-cresol several times and the DNA precipitated with ethanol. 15 μg of DNA was subjected to complete digestion by Hpa 1 according to manufacturer's recommendations (New England Biolabs, Beverly, Mass.) followed by electrophoresis in 0.9% agarose and Southern blotting (12). The nitrocellulose filters were hybridized to a ^{32}P-labeled plasmid containing a 4.4-kb insert of human DNA including the complete β-globin gene (kindly provided by Dr. Arthur Bank) and analyzed, following autoradiography, for the Hpa 1-β-globin gene arrangement as previously described (6, 7) (Fig. 1).

RESULTS

The Togolese population consisted of 20 normals (AA), 3 homozygotes for hemoglobin C (CC), 2 homozygotes for hemoglobin S (SS), 7 double heterozygotes for hemoglobin S-hemoglobin C (SC), 12 hemoglobin S heterozygotes (AS), 6 hemoglobin C heterozygotes (AC), and 1 hemoglobin S/hemoglobin Lepore double heterozygote (Table I). 7 of 41 β^A-genes with certain assignment (40 β^A-genes from normals and 1 β^A-gene from an AC heterozygote with a 13/13 genotype), or 17%, were carried on 13-kb Hpa 1 fragments. The rest were carried on 7.6-kb Hpa 1 fragments. All 12 assignable β^S-genes (4 β^S-genes from SS homozygotes, 7 from SC double heterozygotes, and 1 from the S/

TABLE I
Hpa 1 Genotypes—Togo

Hemoglobin	No.	7.6/7.6	13/13	13/7.6	β^{A*} 13	β^{A*} 7.6	β^S 13	β^S 7.6	β^C 13	β^C 7.6
AA	20	14	0	6	6	34				
SS	2	0	2	0			4			
CC	3	0	3	0					6	
SC	7	0	7	0			7		7	
AS	12	0	0	12						
AC	5	0	1	4	1				1	
S Lepore	1	0	1	0			1			
					17%–13		100%–13		100%–13	

* The columns beneath the headings β^A, β^S, and β^C contain the assignment of the globin genes to either a 13-kb or 7.6-kb Hpa 1 fragment. Thus, 7/41 β^A-genes, or 17%, reside on 13-kb Hpa 1 fragments, whereas 12/12 β^Sgenes and 14/14 β^C-genes, or 100%, reside on 13-kb Hpa 1 fragments.

TABLE II
Hpa 1 Genotypes—Ivory Coast

Hemoglobin	No.	7.6/7.6	13/13	13/7.6	β^{A*} 13	β^{A*} 7.6	β^{S} 13	β^{S} 7.6	β^{C} 13	β^{C} 7.6
AA	31	23	1	7	9	53				
SS	13	8	1	4			6	20		
S Thal	1	0	1	0	1		1			
SC	3	0	0	3						
AS	3	1	1	1	1	1	1	1		
AC	3	0	1	2	1				1	
					18%–13		28%–13		(100%–13)	

* The columns beneath the headings β^{A}, β^{S}, and β^{C} contain the assignment of the globin genes to either a 13-kb or 7.6-kb Hpa I fragment. Thus, 12/66 β^{A} genes, or 18%, reside on 13-kb Hpa 1 fragments, whereas 8/29 β^{S} genes, or 28%, reside on 13-kb Hpa 1 fragments. Only one β^{C}-gene is assignable; its 100% linkage to a 13-kb Hpa 1 fragment is in parentheses.

Lepore double heterozyote) were carried on 13-kb Hpa 1 fragments as were all 14 assignable β^{C}-genes (6 from CC homozygotes, 7 from SC double heterozygotes, and 1 from an AC heterozygote with a 13/13 genotype). Additionally, the S/Lepore double heterozygote has a 13/13 genotype, thus assigning the Lepore gene to a 13-kb Hpa 1 fragment. The other β^{A}-, β^{S}- and β^{C}-genes could not be assigned with assurance to a Hpa 1 fragment.

The Ivory Coast group consisted of 31 normals, 13 SS homozygotes, 1 sickle-β-thalassemia double heterozygote (S/Thal), 3 double SC heterozygotes, 3 AS heterozygotes, and 3 AC heterozygotes (Table II). 12 of 66 β^{A}-genes (62 from normals, 1 AC and 1 AS each with a 13/13 genotype, 1 AS with a 7.6/7.6 genotype, and 1 S/Thal with a 13/13 genotype), or 18%, were linked to the 13-kb Hpa 1 fragment, a result similar to the Togolese β^{A}-gene. By contrast, 21 of 29 assignable β^{S}-genes (26 from SS homozygotes, 1 from an S/Thal with a 13/13 genotype, and 2 from AS heterozygotes with either a 13/13 or 7.6/7.6 genotype), or 72%, were carried on a 7.6-kb Hpa 1 fragment. Only one β^{C}-gene could be assigned with certainty and that was to a 13-kb Hpa 1 fragment.

DISCUSSION

The results demonstrate that the 13-kb Hpa 1 fragment is common in normal subjects from ethnically diverse regions of West Africa with an incidence appreciably higher than that found in the United States (17–18% vs. 3–8%). The lower frequency in Black Americans cannot be explained by admixture with Caucasian genes because the estimate of 22% Caucasian ancestry in Black Americans (13) results in a corrected incidence of only 10–12% at the upper limit. However, the lower incidence in the United States may, at least in part, reflect the diverse geographical and ethnic origins of the slave trade.

The common occurrence of the 13-kb Hpa 1 fragment among normal subjects in West Africa suggests that this DNA alteration is nondeleterious, and that the mutation occurred in an ancient population providing a common ancestry to the presently diverse ethnic groups studied.

By contrast, the β^{S}-Hpa 1 linkage is not homogeneous in West Africa, as previously suggested by Kan and Dozy (9) (Fig. 2). In Togo, the β^{S}-gene is tightly linked to the 13-kb Hpa 1 fragment with an incidence approaching 100% (<1/11), whereas in the Ivory Coast ~72% of the β^{S}-genes are found on 7.6-kb Hpa 1 fragments. Our results suggest that at least two separate chromosomes, one bearing a β^{S}-gene-7.6-kb Hpa 1 fragment, and one bearing a β^{S}-gene-13-kb Hpa 1 fragment, were selected for and expanded in ethnically diverse regions within West Africa. Two alternative mechanisms may explain the data.

(*a*) The β^{S}-gene mutation may have occurred more than once within West Africa, arising at least once on a 13-kb Hpa 1 fragment and at least once on a 7.6-kb Hpa 1 fragment. Under the same selective pressure, two or more sickle mutations were amplified.

(*b*) Recombination between the sites of the β^{S}-gene mutation and the Hpa 1 polymorphism would require the appearance of only one β^{S}-gene mutation. However, studies of the 13-kb Hpa 1 fragment in various populations support the hypothesis that recombination is unlikely to explain the results (14). For example, in Togo, the β^{C}-gene has been found to reside on the 13-kb Hpa 1 fragment in all cases, with a frequency, therefore, of at least 93%. Also, in North Africa the β^{S}-gene is found on the 13-kb Hpa 1 fragment in all cases studied with a frequency of at least 98% (15). In addition, we and others (9) have not observed any crossover events in

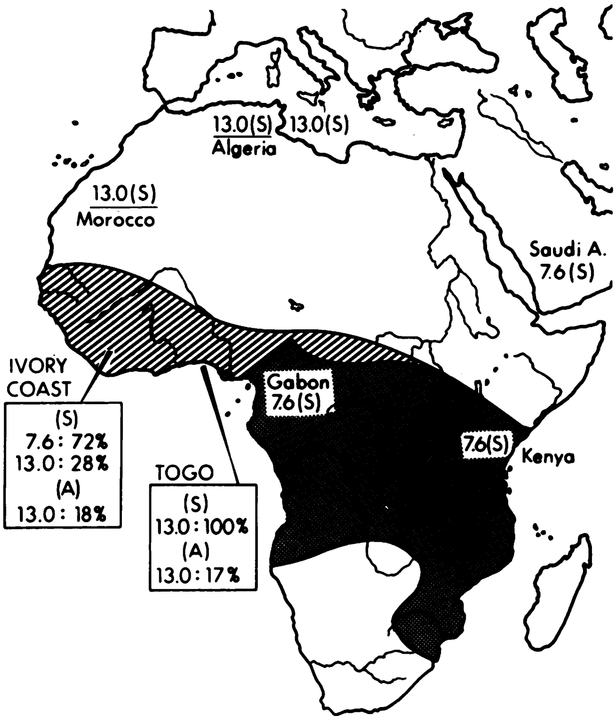

FIGURE 2 Map of Africa with distribution of the Hpa 1-β-globin gene polymorphism. Hatched region represents areas presently inhabited by non-Bantu-speaking Negroid people, whereas the shaded region represents the present extent of Bantu-speaking people (18). Results from Togo include tight linkage of the β^S-gene (S) to the 13-kb Hpa 1 fragment, and a 17% frequency of the β^A-gene (A) residing on the 13-kb Hpa 1 fragment. Ivory Coast data include a similar frequency 18% for the A gene, whereas the S gene is found on a 7.6-kb Hpa 1 fragment 72% of the time. We have previously reported data for North Africa (underlined) (15), and Kan and Dozy have reported results for Gabon, Kenya, Saudi Arabia, and Sicily (9).

family studies. Finally, if recombination occurred between a β^S-7.6-kb Hpa 1 fragment and a β^A-13-kb Hpa 1 fragment one should expect the expansion of both genes, as no evidence exists of a differential value for selection (6, 7). The tight linkage, in Togo, of the β^S-gene to the 13-kb Hpa 1 fragment therefore argues against recombination.

For these reasons, we favor the first explanation, although formal proof will require the identification of other polymorphisms lying between the sickle mutation and the Hpa 1 polymorphic site. Such polymorphisms are being sought but have not yet been identified.

Other polymorphisms in linkage disequilibrium with the β^S-gene have been described (16). However, the Hind III polymorphisms located in the large introns of the $^G\gamma$- and $^A\gamma$-genes lie ~22 kb 5′ to the β-gene locus at a minimum. Analysis of these polymorphisms will not strengthen the argument for or against independent mutations vs. recombination, because they occur at a distance from the locus of interest, and since chromosomes bearing the β^S-gene-7.6-kb Hpa 1 fragment have been shown to possess several different Hind III polymorphism patterns (16). Because the available evidence supports a low incidence of recombination, the linkage heterogeneity in the Ivory Coast may reflect the appearance of the independent β^S-gene mutation on a 7.6-kb Hpa 1 bearing chromosome as well as the more recent diffusion of the β^S-gene-13-kb Hpa 1 chromosome from the Togo-Upper Volta region.

Interestingly, a patient with hemoglobin Lepore (17) was studied, and the Lepore gene was found to reside on a 13-kb Hpa 1 fragment, which is expected, if the crossover involved a β-globin gene on a 13-kb Hpa 1 fragment. One patient with hemoglobin Lepore has been studied by Kan and Dozy (6) and found to possess a 7.3-kb Hpa 1 fragment. Theoretically, our patient should possess a 12.7-kb fragment rather than the observed 13-kb fragment, but agarose gel electrophoresis cannot discriminate between these size fragments. Since the fragment is not appreciably shortened, the fusion gene does not contain the Hpa 1 site found within the large intron of the normal δ-gene demonstrating that in this individual the crossover point is 5′ to the site. This patient provides an opportunity for studying the diffusion of a Lepore mutation to other regions.

Diffusion of the β^S gene from West Africa, while under selective pressure from malaria, may not be accompanied by the β^A-13-kb Hpa 1 chromosome, assuming the 13-kb Hpa 1 genotype per se does not provide a selective advantage. We have previously demonstrated that the β^S-gene found in North African Arab countries is found tightly linked to the 13-kb Hpa 1 fragment, whereas the β^A-gene is tightly linked to the 7.6-kb Hpa 1 fragment (15). The evidence suggests that the North African β^S-gene diffused from the Togo-Upper Volta region rather than from the Middle East where the β^S-gene is commonly found on the 7.6-kb Hpa 1 fragment (8) (Fig. 2).

Anthropologic data, including studies of language structure, suggest that the Bantu-speaking peoples of Equatorial and East Africa may be derived from an ancient West African group, most likely located in margins of the Benue River Valley, which is in eastern Nigeria near the border with Cameroun (18). A small number of samples from Gabon and Kenya (inhabited by Bantu-speaking people) demonstrate the linkage of the β^S-gene to the 7.6-kb Hpa 1 fragment (9) (Fig. 2). It would be of interest to study individuals from the Benue Valley and the prediction would be, if the linguistic analysis is correct, that this region contains individuals with the β^S-gene-7.6-kb Hpa 1 fragment. The problem is what connection exists, if any, between the originators of the great Bantu migration and the presence of the β^S-gene-7.6-kb Hpa 1 fragment in the Ivory Coast. Only further data can resolve this question. The presence of the β^S-gene-7.6-kb Hpa 1 fragment in Arabia and India is compatible with the well-known commercial and slave trade routes between East Africa and the Arabian peninsula and the Indian subcontinent.

The Hpa 1 polymorphism can be employed in prenatal diagnosis of sickle-cell anemia in selected populations (16, 19). The tight β^S-13-kb Hpa 1 linkage is exploitable in North Africa because the β^A-gene is tightly linked to the 7.6-kb Hpa 1 fragment (15). This is not true for Togo, where the β^A-gene is commonly found on the 13-kb Hpa 1 fragment. Thus, this observation, and the heterogeneous distribution of the β^S- and β^A-gene linkage in the Ivory Coast, make prenatal diagnosis solely by Hpa 1 analysis for people derived from these regions unreliable. Studies of other linked polymorphisms may provide in the future better gene assignment, thus making prenatal diagnosis by fetal DNA analysis widely applicable.

These studies demonstrate that restriction endonuclease analysis of linked DNA polymorphisms is a powerful tool for the study of the origin and evolution of a structural gene mutation. The increasing availability of probes for other human structural genes will provide means to study linked DNA polymorphisms for a widening variety of genetic disorders. Additionally, this approach may provide an avenue both for the prenatal diagnosis of other genetic disorders, and for studies directed to the understanding of the wide clinical variability often associated with an apparently single structural gene mutation.

ACKNOWLEDGMENTS

We thank Professor Keke, Dean of the Faculty of Medicine, Universite du Benin, Lome, Togo, as well as the Faculty of Medicine, University d'Abidjan, Abidjan, Ivory Coast for their support.

This work was supported in part by National Institutes of Health grants GM 28841, HL 21016, Sinsheimer Foundation, and Institut National de la Santé et de la Recherche Médicale (France) 79.5.170.3.

REFERENCES

1. Bunn, H. F., B. F. Forget, and H. M. Ranney. 1977. Human Hemoglobins. W. B. Saunders Company, Philadelphia, Pa.
2. Cavalli-Sforza, L. L., and W. F. Bodmer. 1971. The Genetics of Human Populations. W. H. Freeman & Company Publishers, San Francisco, Calif.
3. Mears, J. G., F. Ramirez, D. Leibowitz, and A. Bank. 1978. Organization of human δ and β globin genes in cellular DNA and the presence of intragenic inserts. *Cell.* **15:** 15–23.
4. Flavell, R. A., J. M. Kooter, E. De Boer, P. F. R. Little and R. Williamson. 1978. Analysis of the $\beta\delta$ globin gene loci in normal and Hb Lepore DNA: Direct determination of gene linkage and intergene distance. *Cell.* **15:** 25–41.
5. Bank, A., J. G. Mears and F. Ramirez. 1980. Disorders of human hemoglobin. *Science (Wash., D. C.).* **207:** 486–493.
6. Kan, Y. W., and A. M. Dozy. 1978. Polymorphism of DNA sequence adjacent to human β globin structural gene: relationship to sickle mutation. *Proc. Natl. Acad. Sci. U.S.A.* **75:** 5631–5635.
7. Feldenzer, J., J. G. Mears, A. L. Burns, C. Natta, and A. Bank. 1979. Heterogeneity of DNA fragments associated with the sickle-globin gene. *J. Clin. Invest.* **64:** 751–755.
8. Panny, S. R., A. F. Scott, K. D. Smith, J. A. Phillips III, H. H. Kazazian, Jr., C. C. Talbot, Jr., and C. D. Boehm. 1981. Population heterogeneity of the Hpa 1 restriction site associated with the β globin gene: Implications for prenatal diagnosis. *Am. J. Hum. Genet.* **33:** 25–35.
9. Kan, Y. W., and A. M. Dozy. 1980. Evolution of the hemoglobin S and C genes in world populations. *Science (Wash. D. C.).* **209:** 388–391.
10. Livingstone, F. B. 1967. Abnormal Hemoglobins in Human Populations. Aldine Publishing Company, Chicago, Ill. 76–84.
11. Wajcman, H., K. P. E. Amegnizin, O. Belkhodja, and D. Labie. 1977. Hemoglobin J Lome β59 (E3) Lys $\rightarrow$ Asn. A new fast moving variant found in a Togolese. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* **84:** 372–374.
12. Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98:** 503–517.
13. Reed, T. E. 1969. Caucasian genes in American negroes. *Science (Wash. D. C.).* **165:** 762–768.
14. Kurnit, D. M. 1979. Evolution of sickle gene variant. *Lancet.* **I:** 104.
15. Mears, J. G., C. Beldjord, M. Benabadji, Y. A. Belghiti, M. A. Baddou, D. Labie, and R. L. Nagel. 1981. The sickle gene polymorphism in North Africa. *Blood.* In press.
16. Phillips, J. A., S. R. Panny, H. H. Kazazian, Jr., C. D. Boehm, A. F. Scott, and K. D. Smith. 1980. Prenatal diagnosis of sickle cell anemia by restriction endonuclease analysis: Hind III polymorphisms in γ-globin genes extend test applicability. *Proc. Natl. Acad. Sci. U. S. A.* **77:** 2853–2856.
17. Amegnizin, K. P. E. 1980. Hemoglobin Lepore Boston au Togo. *Med. d'Afr. Noire.* **27:** 437–440.
18. Greenberg, J. H. 1973. African languages. *In:* Peoples and Culture of Africa. E. P. Skinner, editor. Doubleday & Co., Inc. New York. 71–80.
19. Kan, Y. W., and A. M. Dozy. 1978. Antenatal diagnosis of sickle-cell anaemia by DNA analysis of amniotic-fluid cells. *Lancet.* **II:** 910–912.